Technical note

Conceptual analogy between mass density and probability density – An almost non-technical note

László Bokor*

Budapest, Hungary

* Correspondence: Email: bokorl@proton.me.

Academic Editor: Chris Tisdell

Abstract: I show how multivariate integration over a physical object serves as a handy tool for a gentle introduction to multivariate probability theory with continuous variables. This exercise helps to visualise and thus link concepts that at first sight seem distant or even unrelated. The concepts and methods presented are digestible for advanced high school classes and above.

Keywords: multivariate probability, continuous random variable, density, physical space, analogy, dimensionality, shape

1. Introduction

Integration can be used to derive the mass of non-homogeneous objects if the structure is not too complex. Such an exercise, however, is not only useful in engineering, but can also be a very illustrative, tangible way of introducing multivariate probability theory with continuous variables. It seems logical to start with the three-dimensional case, since this is what we encounter every day, and thus what we can easily relate to.

2. Mass of a physical object

Take a non-homogeneous cube with an edge of a meters. Suppose that its density (ρ) depends on the sum of the shortest distances from the three faces meeting at a given vertex. Thus, placing the cube into a Cartesian coordinate system where the origin is that particular vertex (Figure 1), the

density can be expressed as the sum of the spatial coordinates,[1]

$$\rho(x, y, z) = (x + y + z)\frac{\text{kg}}{\text{m}^4}.$$
(1)

It follows that the density of the large cube is a spectrum between 0 kg/m^3 and $3a$ kg/m^3. Since the equation $x + y + z = j \in \mathbb{R}^+$ m defines a hyperplane, the cube consists of a continuum of parallel cross-sections (slices). Practically speaking, each slice is a set of points for which the sums of the coordinates are equal (Figure 2).
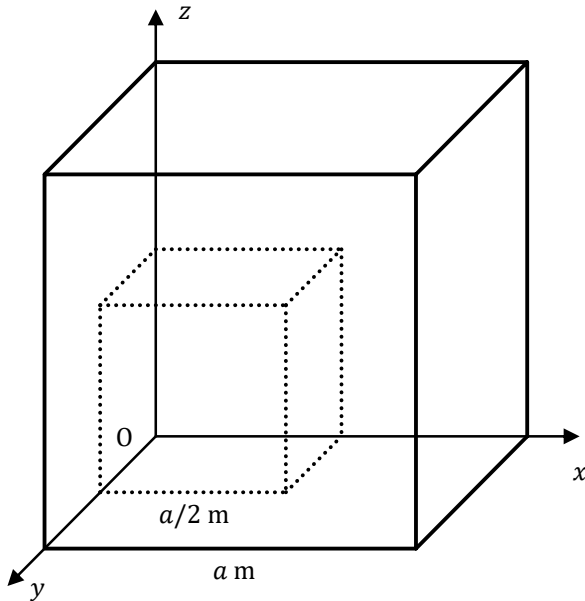

Figure 1. Perspective view in a Cartesian coordinate system.
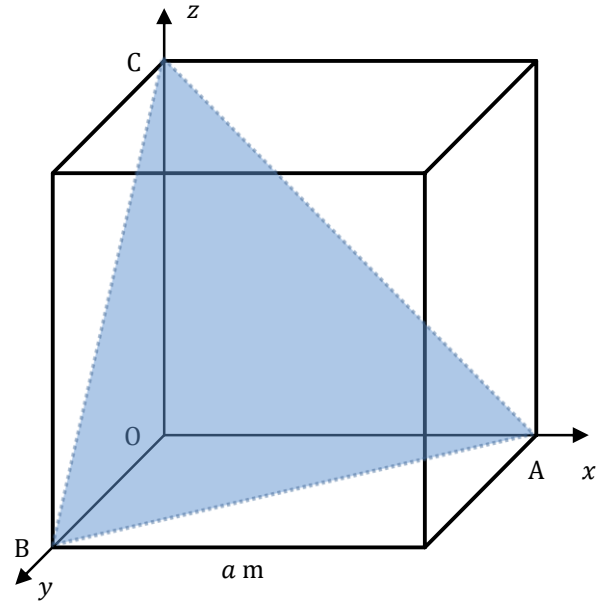
Figure 2. Cross-section ABC at $j = a$.

Question 1: *What is the mass of the small cube with an edge of $a/2$ meters and its relation to the mass of the large cube?*

If the material were homogeneous, the answer would be straightforward. The volume of the small cube is $a^3/8$ m^3, which should be multiplied by the density of the material to get the mass. In our case, however, the density is in not a constant but a linear function of x, y, z. Triple integration comes to our rescue.

In this process, the object is scanned properly along each dimension. Since we are dealing with cubes that have three edges on the axes of the coordinate system, we have an easy job: The limits of integration, i.e. the lower and upper bounds of the scan are simply the origin and the edge length in each dimension. That is, the mass of the small cube is (substituting Equation 1)[2]

---

[1] Note that the unit of measurement of the variables is meter, and so $\text{m} \cdot \text{kg/m}^4 = \text{kg/m}^3$.

[2] Only power rule is applied: $\int x^n \, dx = x^{n+1}/(n + 1) + c$. Note that the constant cancels out for a definite integral. For simplicity, the units of measurement were omitted during the integration. However, notice that, at the end, the meters would be to the fourth power, which together with the factor kg/m^4 of the density function is $\text{m}^4 \cdot \text{kg/m}^4 = \text{kg}$.

$$\int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} \rho(x,y,z)\, dz\, dy\, dx = \int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} (x+y+z)\, dz\, dy\, dx$$

$$= \int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} \left[xz + yz + \frac{z^2}{2} \right]_0^{\frac{a}{2}} dy\, dx = \int_0^{\frac{a}{2}} \int_0^{\frac{a}{2}} \left(\frac{a}{2}x + \frac{a}{2}y + \frac{a^2}{8} \right) dy\, dx$$

$$= \int_0^{\frac{a}{2}} \left[\frac{a}{2}xy + \frac{a}{4}y^2 + \frac{a^2}{8}y \right]_0^{\frac{a}{2}} dx = \int_0^{\frac{a}{2}} \left(\frac{a^2}{4}x + \frac{a^3}{8} \right) dx = \left[\frac{a^2}{8}x^2 + \frac{a^3}{8}x \right]_0^{\frac{a}{2}} \tag{2}$$

$$= \frac{a^4}{32} + \frac{a^4}{16} = \frac{3a^4}{32} \text{ kg}.$$

The mass is calculated in the same way for the large cube:

$$\int_0^a \int_0^a \int_0^a \rho(x,y,z)\, dz\, dy\, dx = \int_0^a \int_0^a \int_0^a (x+y+z)\, dz\, dy\, dx$$

$$= \int_0^a \int_0^a \left[xz + yz + \frac{z^2}{2} \right]_0^a dy\, dx = \int_0^a \int_0^a \left(ax + ay + \frac{a^2}{2} \right) dy\, dx$$

$$= \int_0^a \left[axy + \frac{a}{2}y^2 + \frac{a^2}{2}y \right]_0^a dx = \int_0^a (a^2x + a^3)\, dx = \left[\frac{a^2}{2}x^2 + a^3x \right]_0^a \tag{3}$$

$$= \frac{a^4}{2} + a^4 = \frac{3a^4}{2} \text{ kg}.$$

Table 1 summarises our answer to Question 1.

Table 1. Masses and their ratios.

	Small cube	Large cube	Ratio
Homogeneous*	$\dfrac{3a^4}{8}$ kg	$3a^4$ kg	$\dfrac{1}{8}$
Non-homogeneous	$\dfrac{3a^4}{32}$ kg	$\dfrac{3a^4}{2}$ kg	$\dfrac{1}{16}$

*$\rho = 3a$ kg/m^3 (i.e., the spectral maximum of the non-homogenous) is assumed.

Of a homogeneous object and its non-homogenous lookalike, the one with the lower average density is the lighter. Since the density of the homogeneous is assumed to be equal to the spectral maximum of the non-homogeneous, it is obvious why the homogeneous is the heavier.

Overall, for complex structures, it is usually necessary to go beyond simple geometric formulae and use integration to calculate the mass. Depending on the shape of the surface and also the density

structure, other coordinate systems may be more useful. For example, in the spherical coordinate system, the three coordinates of a point P are the radial distance (i.e., the length of the line segment OP), the azimuthal angle (measured between the x-axis and OP's projected line segment on the xy-plane) and the zenith angle (measured between the z-axis and OP). With a concrete example, the point (a, a, a) in our Cartesian coordinate system would be the point $(a\sqrt{3}, \pi/4, \pi/4)$ in a spherical coordinate system assuming identical axes, where the two $45°$ angles are in radians. Without going into the details, this alone suggests that for a non-homogeneous sphere, the radial and angular scans under the spherical coordinate system (with more obvious limits of integration) may be more appropriate than the scans along the axes under the Cartesian coordinate system.

3. From mass density through occurrence density to …

The above concept can be readily applied to probability calculations. Let us suppose that a source emits sinusoidal pulses for which all their characteristics (amplitude, wavelength, speed) are random variables. Assume that each variable can take a value between 0 and a in their own unit of measurement, and that a pulse is more likely to occur if the value of any of its properties is greater.[3] Assuming direct proportionality with proportionality constant 1 for each variable and, for simplicity, taking the values without the units of measurement, it means

$$f(x, y, z) = x + y + z \,, \tag{4}$$

where x, y and z are the values for amplitude, wavelength and speed, respectively.[4] Thus, we know the ranges (0 to a for each), we know the occurrence density of the triplets (defined by function f), and so we know the set of all possible outcomes of our random experiment, i.e., the sample space (S).[5] This abstract space can be imagined as if it were a real space. That is, x, y, and z can be thought of as "spatial" coordinates and the function f as the descriptor of "mass" density.

Question 2: *What is the probability that the next pulse emitted will have amplitude and wavelength of $a/2$ meters or less and speed of $a/2$ m/s or less?*

Actually, we have already solved this puzzle. The set of outcomes defined in Question 2 is a so-called event (E), therefore the probability in question is (recall Table 1)

$$Pr(E) = \frac{E}{S} = \frac{\text{"Mass" of small cube}}{\text{"Mass" of large cube}} = \frac{\frac{3a^4}{32}}{\frac{3a^4}{2}} = \frac{1}{16} \,, \tag{5}$$

where Pr is a function called probability measure that maps all events to a real number between 0 and 1, i.e., $Pr: Event \rightarrow [0,1]$. That is, the domain of this function is a collection of sets.

For example, let us take two events whose union gives the sample space: The one that we have already defined (E), and its complement set (E^C), where all three variables are greater than $a/2$ (in

---

[3] This is just an arbitrary example with no underlying physical content.

[4] If the function were defined for the values with units of measurement, it would be $f(x, y, z) = x/\text{m} + y/\text{m} + z \cdot \text{s}/\text{m}$ similar to Equation 1.

[5] Note that, in a continuous space, a single point (triplet) has no mass and hence no density, but the very small environment around it has mass and hence density.

their own unit of measurement). Then, the domain consists of these two disjoint events which the probability measure maps to 1/16 and 15/16 (recall Equation 5), plus the trivial events, the empty set ($\emptyset$) and the sample space (S), which it maps to 0 and 1.[6] Note that non-trivial events are not necessarily defined as disjoints. The point is that all outcomes should be covered, i.e., the union of non-trivial events should be the sample space. Empty set can be, for example, an elementary event, meaning an event having only a single outcome (e.g., "Each variable equals $a/2$"). By geometric analogy, a point has no extent and hence no mass, i.e. the numerator in Equation 5 and so the probability will be zero ($Pr: \emptyset \rightarrow 0$).

4.   … probability density

Recall that the "mass" of the large cube is $3a^4/2$. Thus, with $a := \sqrt[4]{2/3}$, our non-homogenous structure of the large cube represents not only an occurrence density, but also a probability density: It integrates to 1 just like the probability density function (PDF) of a normal (Gaussian) or any other absolutely continuous probability distribution.[7] Practically, as the "mass" of the sample space is 1, it means that the "mass" of an event is also the probability of the event (e.g., $\Pr(E) = "Mass"\ of\ small\ cube/1$).

Thus, in this case, Equation 4 is a three-variable joint PDF. A three-variable PDF has three marginal PDFs. The marginal PDF shows the probability distribution of a variable considering all related value pairs of the remaining two variables. For example, it shows the probability distribution of amplitude considering all related pairs of wavelength and speed. Formally, the marginal PDF of amplitude is (using the expression $a^2x + a^3$ obtained in Equation 3 after the integration with respect to z and y, and substituting $a = \sqrt[4]{2/3}$)

$$f_{AMP}(x) = x\sqrt{\frac{2}{3}} + \sqrt[4]{\left(\frac{2}{3}\right)^3}.$$ (6)

Question 3: *What is the probability that the next pulse emitted will have amplitude of a/2 meters or less?*

To answer this question, we need to integrate $f_{AMP}(x)$ from 0 to $\sqrt[4]{2/3}/2$ (i.e., halfway). Visually, the integration of this function can be also thought of as an underlying cross-section scan of the large cube along the desired segment of the x-axis. That is, the "mass" of the ABCDEFGO cuboid in Figure 3 and the size of the shaded area in Figure 4 are exactly the same, namely the probability we want to find.

---

[6] This collection $\{\emptyset, E, E^C, S\}$ is a so-called σ-algebra on S. The trio of sample space, σ-algebra, and probability measure is called the probability space.

[7] The Cantor distribution is a notable example of a continuous (but not absolutely continuous) distribution without a PDF. For a geometric discussion of this interesting distribution, see Presnell [1].
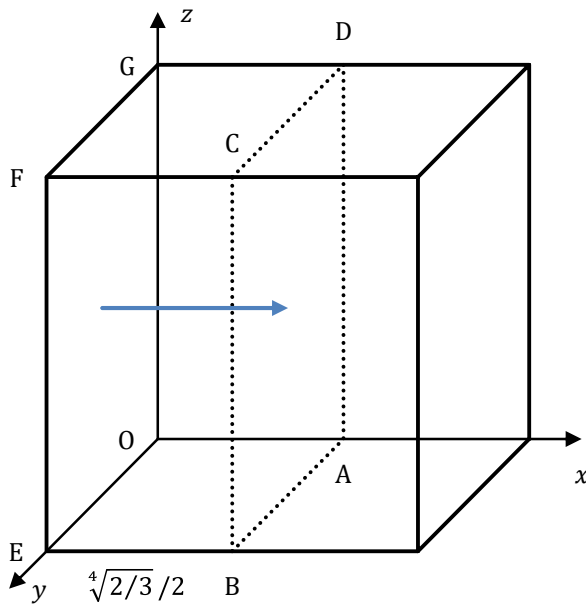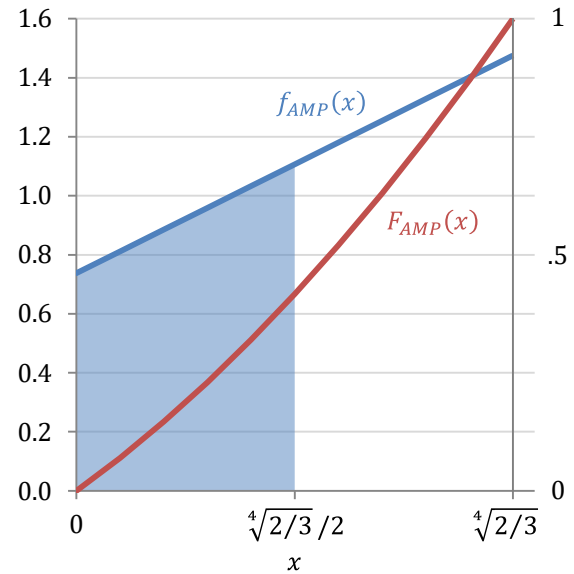
Figure 3. Cross-section scan along x.

Figure 4. Marginal PDF and CDF (r. axis) of amplitude.

The cumulative distribution function (CDF) simply shows the integrated "mass" of the PDF from the minimum value of the random variable to a specific value of the random variable, i.e., the probability that the random variable is smaller than the specific value. As for amplitude, the event is $amplitude \leq x$, and so the related probability is $Pr(amplitude \leq x)$. It logically follows that the CDF goes from 0 to 1 as x increases from its minimum value. Formally, the marginal CDF of amplitude is (using Equation 6)

$$F_{AMP}(x) = \int_0^x f_{AMP}(t)\,dt = \int_0^x \left(t\sqrt{\frac{2}{3}} + \sqrt[4]{\left(\frac{2}{3}\right)^3} \right) dt = \frac{x^2}{2}\sqrt{\frac{2}{3}} + x\sqrt[4]{\left(\frac{2}{3}\right)^3}$$

$$\equiv F\left(\sqrt[4]{\frac{2}{3}}, \sqrt[4]{\frac{2}{3}}, x \right) = \int_0^{\sqrt[4]{\frac{2}{3}}} \int_0^{\sqrt[4]{\frac{2}{3}}} \int_0^x f(t, y, z)\,dz\,dy\,dt ,$$

(7)

where the expressions after sign $\equiv$ recalls the underlying content highlighted in Figure 3. The specific value from Question 3 is $x = \sqrt[4]{2/3}/2$, so substituting this into Equation 7 gives the probability sought.

To conclude this section, let us invoke our old acquaintance, the small cube. The probability of the event it represents (here: "All variables are $\sqrt[4]{2/3}/2$ or less in their own unit of measurement") is, of course, already known. Notice, however, that since it is located at the origin (i.e., at the minimum values of the variables), the probability can be expressed by the joint CDF,

$$F\left(\frac{\sqrt[4]{\frac{2}{3}}}{2}, \frac{\sqrt[4]{\frac{2}{3}}}{2}, \frac{\sqrt[4]{\frac{2}{3}}}{2}\right) = \frac{1}{16}. \tag{8}$$

5. Towards generalisation with a bit of formalism

Let us return to the general concept of "small cube / large cube" with $a > 0$. The concept can be easily generalized to higher (or lower) dimensions. For example, if the duration of a pulse varies, this new dimension can be added to our sample space. Assuming a range of 0 to a seconds, the sample space becomes a tesseract, i.e., a 4-dimensional hypercube.

Let w, x, y, and z denote the values (without the units of measurement) for duration, amplitude, wavelength and speed, respectively. Suppose that, in the occurrence density function, the duration also has the same proportionality as the other addends. Thus, the sample space is (denoting dimension in parenthesis)

$$S^{(4)} = \left\{ v = (w, x, y, z) \in \mathbb{R}^4 \left| \begin{matrix} 0 < w \le a \\ 0 < x \le a \\ 0 < y \le a \\ 0 < z \le a \end{matrix} \right. \right\}, \tag{9}$$

whereas the occurrence density function is $f^{(4)}(w, x, y, z) = w + x + y + z$. Using the result of Equation 3, the "mass" of such an object is

$$\iiint\int_{S^{(4)}} f^{(4)}(w, x, y, z)\, dz\, dy\, dx\, dw = \int_0^a \int_0^a \int_0^a \int_0^a (w + x + y + z)\, dz\, dy\, dx\, dw$$

$$= \int_0^a \frac{3a^4}{2}\, dw + \int_0^a \int_0^a \int_0^a \int_0^a w\, dz\, dy\, dx\, dw = \left[\frac{3a^4}{2} w\right]_0^a + \left[\frac{a^3}{2} w^2\right]_0^a \tag{10}$$

$$= \frac{3a^5}{2} + \frac{a^5}{2} = 2a^5.$$

The event in which the duration is $a/2$ seconds or less, in addition to the other three analogous conditions, is

$$E^{(4)} = \left\{ v = (w, x, y, z) \in \mathbb{R}^4 \left| \begin{matrix} 0 < w \le a/2 \\ 0 < x \le a/2 \\ 0 < y \le a/2 \\ 0 < z \le a/2 \end{matrix} \right. \right\}, \tag{11}$$

and so its "mass", using the result of Equation 2, is

$$\iiint\int_{E^{(4)}} f^{(4)}(w,x,y,z)\,dz\,dy\,dx\,dw = \int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}(w+x+y+z)\,dz\,dy\,dx\,dw$$

$$= \int_0^{\frac{a}{2}}\frac{3a^4}{32}\,dw + \int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}\int_0^{\frac{a}{2}}w\,dz\,dy\,dx\,dw = \left[\frac{3a^4}{32}w\right]_0^{\frac{a}{2}} + \left[\frac{a^3}{16}w^2\right]_0^{\frac{a}{2}} \qquad (12)$$

$$= \frac{3a^5}{64} + \frac{a^5}{64} = \frac{a^5}{16},$$

and so the probability of such an event is

$$\Pr\!\left(E^{(4)}\right) = \frac{E^{(4)}}{S^{(4)}} = \frac{\frac{a^5}{16}}{2a^5} = \frac{1}{32}. \qquad (13)$$

6. Dimensionality and shape issues

Notice that the probability of our 4-dimensional event (see Equation 13) is only half of the probability of our 3-dimensional event (see Equation 5). It is easy to see that the probabilities of events defined along similar patterns dwarf as the dimension increases. Assuming an occurrence density of 1, for simplicity, the limit is

$$\lim_{n\to\infty}\frac{E^{(n)}}{S^{(n)}} = \lim_{n\to\infty}\frac{\left(\frac{a}{k}\right)^n}{a^n} = \lim_{n\to\infty}\frac{1}{k^n} = 0\,. \qquad (14)$$

where $n \in \mathbb{N}$ denotes the dimension and $k > 1$ is a constant.

Note that, in this context, the shape of the sample space and the event is not irrelevant. Think of a homogeneous spherical sample space and a cubical event with radius (r) and edge (l) of a "metres" and identical center of symmetry. It follows that $E^{(3)}_{cube\ (l=a)} \subset S^{(3)}_{sphere\ (r=a)}$ as the body diagonal of the cube is smaller than the diameter of the sphere $(a\sqrt{3} < 2a)$. As n goes to infinity, however, their ratio also goes to infinity,

$$\lim_{n\to\infty}\frac{E^{(n)}_{cube\ (l=a)}}{S^{(n)}_{sphere\ (r=a)}} = \lim_{n\to\infty}\frac{a^n}{a^n\sqrt{\pi^n}/\Gamma\left(\frac{n}{2}+1\right)} = \lim_{n\to\infty}\frac{\Gamma\left(\frac{n}{2}+1\right)}{\sqrt{\pi^n}} = \infty\ \unicode{x21af}\,, \qquad (15)$$

where Γ denotes the gamma function.[8] That is, beyond a certain point, $E^{(n)}_{cube\ (l=a)} \not\subset S^{(n)}_{sphere\ (r=a)}$ is inevitable, which relation makes no sense.

7. Conclusions

I proposed an approach for those who prefer visualisation to abstraction when they first

---

[8] For an in-depth discussion of the interesting properties of n-sphere's volume (e.g., besides its asymptotic nullity, its non-monotonicity to n), see Smith and Vamanamurthy [2].

encounter a more complex subject such as multivariate probability theory with continuous variables. Once the link is established through a three-dimensional physical example, one could easily move towards abstraction. In fact, geometry as a complement can help to understand several algebraic concepts. Adventurous readers with a deeper interest in geometric probability should delve into the books of, for example, Solomon [3], Klain and Rota [4], or Mathai [5].

Use of AI tools declaration

The author declares that he did not use Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The author thanks the anonymous reviewers for their helpful comments.

Conflict of interest

The author declares that there is no conflict of interest to disclose.

Ethics declaration

The author declares that there was no human participation in this article and therefore ethics approval and informed consent were not required.

References

1. Presnell, B., A Geometric Derivation of the Cantor Distribution. *The American Statistician*, 2022, 76(1): 73–77. https://doi.org/10.1080/00031305.2021.1905062
2. Smith, D.J., Vamanamurthy, M.K., How Small Is a Unit Ball? *Mathematics Magazine*, 1989, 62(2): 101–107. https://doi.org/10.1080/0025570X.1989.11977419
3. Solomon, H., *Geometric Probability – CBMS-NSF Regional Conference Series in Applied Mathematics No. 28*, 1978, Society for Industrial and Applied Mathematics.
4. Klain, D.A., Rota, G-C., *Introduction to Geometric Probability*, 1997, Cambridge University Press.
5. Mathai, A.M., *An Introduction to Geometrical Probability: Distributional Aspects with Applications*, 1999, CRC Press.

Author's biography

Dr. László Bokor is an economist. His research interests are monetary policy, antitrust economics, and green finance.